

In Silico Exploration of the Molecular Mechanism of Clinically Oriented Drugs for Possibly Inhibiting SARS-CoV-2's Main Protease

Tien Huynh, Haoran Wang, and Binqun Luan*



Cite This: *J. Phys. Chem. Lett.* 2020, 11, 4413–4420



Read Online

ACCESS |



Metrics & More

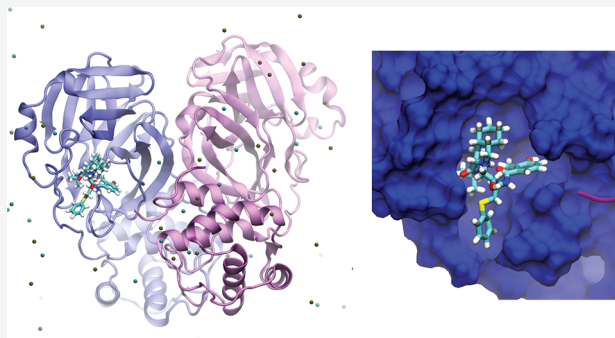


Article Recommendations



Supporting Information

ABSTRACT: Currently, the new coronavirus disease 2019 (COVID-19) is a global pandemic without any well-calibrated treatment. To inactivate the SARS-CoV-2 virus that causes COVID-19, the main protease (Mpro) that performs key biological functions in the virus has been the focus of extensive studies. With the fast-response experimental efforts, the crystal structures of Mpro of the SARS-CoV-2 virus have just become available recently. Herein, we theoretically investigated the mechanism of binding between the Mpro's pocket and various marketed drug molecules being tested in clinics to fight COVID-19 that show promising outcomes. By combining the existing experimental results with our computational ones, we revealed an important ligand binding mechanism of the Mpro, demonstrating that the binding stability of a ligand inside the Mpro pocket can be significantly improved if part of the ligand occupies its so-called "anchor" site. Along with the highly potent drugs and/or molecules (such as nelfinavir) revealed in this study, the newly discovered binding mechanism paves the way for further optimizations and designs of Mpro's inhibitors with a high binding affinity.



Coronavirus disease 2019 (COVID-19) is a viral respiratory disease of zoonotic origin caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus. COVID-19 first emerged in the city of Wuhan (China) at the end of 2019 but now has turned into a global pandemic reported in all continents just after a few short months. SARS-CoV-2 appears to be highly contagious and spreads mainly from human to human through respiratory droplets from coughing and sneezing of the infected persons as well as by fomites. SARS-CoV-2 belongs to a family of viruses named coronaviruses for the crownlike spikes on their surface that can infect bats, birds, pigs, cows, and other mammals and mutate easily to transfer from animals to humans.¹ Before the COVID-19 outbreak, six strains of such virus already have been identified as human pathogens known to cause viral respiratory illness. However, not all of them are highly pathogenic. For examples, HCoV-229E, HCoV-NL63, HCoV-OC43, and HCoV-HKU1 merely cause a common cold. In contrast, both the severe acute respiratory syndrome coronavirus (SARS-CoV)² and Middle East respiratory syndrome coronavirus (MERS-CoV)³ have caused large-scale outbreaks during the past two decades with significant case-fatality rates (9.6% for SARS and 34% for MERS). As for COVID-19, its case-fatality rate remains uncertain given the pandemic is still in its early stages.

Currently, it is well-known that the SARS-CoV-2's main protease (Mpro) constitutes one of the most attractive antiviral drug targets, because the viral maturation almost exclusively

relies on the Mpro's activity. For example, maturation of 12 nonstructural proteins (Nsp4–Nsp16), including critical proteins like the RNA-dependent RNA polymerase (RdRp, Nsp12) and helicase (Nsp13), requires the cleavage through the Mpro. It has been demonstrated in experiment that the Mpro inhibition prevented viral replication in multiple studies.^{4,5} Considered as the Achilles' heels of SARS-CoV-2, the Mpro is therefore among the top candidates for drug discovery. Additionally, the Mpro's inhibitor(s) is likely to inactivate virus in different cell types in different organs, independent of the various receptors/host proteases (on the cell membrane) required for virus entry.

So far, a specific Mpro inhibitor is still missing for the SARS-CoV-2 virus. Irreversible inhibitors like N3 are efficacious and have been proven to inhibit SARS-CoV-2 virus in *in vitro* viral proliferation models with moderate efficacy ($EC_{50} = 4\text{--}5\ \mu\text{M}$).⁵ However, development of these tool drugs into an approved drug could take years to accomplish. In an *BioRxiv* preprint,⁶ a few marketed drug such as ebiselen, disulfiram, tideglusib, and carmofur have exhibited EC_{50} values of 0.67

Received: March 30, 2020

Accepted: May 14, 2020

Published: May 14, 2020



μM , 9.35 μM , 1.55 μM , and 1.82 μM respectively with an *in vitro* enzymatic assay, which translate to an EC_{50} of 4.6 μM in antiviral activity for ebselen (best in class), compared to an EC_{50} of 16.77 μM for N3.⁵ These experiments validated that Mpro could be a viable antiviral target, albeit additional efforts are needed to search for more potent and specific antiviral drugs with a better safety margin than ebselen that is an (irreversible) inhibitor for the Mpro and many other enzymes in a broad spectrum of tissues with significant cellular toxicity.⁷ Motivated by the fact that Mpro can be inhibited by multiple drug-like ligands, we speculated that a range of drug molecules may efficaciously interact with the Mpro pocket. Given the urgency, we used *in silico* methods to explore a set of 19 marketed drugs that have exhibited a great deal of promise in clinics, aiming to identify the potential high-potential ones for the Mpro inhibition and discover a common binding mechanism for these drug molecules inside the Mpro's pocket.

Understanding the structural determinants for protein–ligand complex at the atomic level is crucial for designing ligands with high specificity and affinity for a target protein. Moreover, gaining insight into the mechanisms responsible for the protein–ligand recognition and binding greatly facilitates the discovery and development of drugs for the treatment of the underlying disease. We carried out all-atom molecular dynamics (MD) simulations that are widely used in the studies of biomolecules,^{8,9} guided with fast and efficient docking studies. Besides the identification of several high-potency drugs and/or molecules, we unveiled the consensus binding mechanism that a ligand prefers to bind the “anchor” site of the Mpro pocket, which might facilitate the future design and optimization of an inhibitor for the SARS-CoV-2's Mpro.

In our *in silico* studies, we used the NAMD¹⁰ package for studying the structures of the stand-alone apo Mpro as well as the ligand-bound one. We first equilibrated the crystal structure of the SARS-CoV-2 Mpro [Protein Data Bank (PDB) entry 6LU7] in the physiologically relevant environment. The simulation system is illustrated in Figure 1. The dimer structure of the Mpro (colored blue and purple) is illustrated in the cartoon representation. The entire protein was then solvated in a water box with dimensions of 97.4 Å \times 97.4 Å \times 97.4 Å. Eighty-eight K^+ and 80 Cl^- ions were added to the solution to neutralize the net charge of the protein and set the ion concentration to 0.15 M. The MD simulations were carried out on the IBM Power-cluster. We applied the TIP3P model^{11,12} for water, the standard ion force field,¹³ and the CHARMM36 force field¹⁴ for the protein. A smooth cutoff (10–12 Å) was used to calculate van der Waals energies. Electrostatic interactions were calculated using the particle-mesh Ewald (PME) method (grid size of ~ 1 Å). With the SETTLE algorithm¹⁵ enabled to keep all bonds rigid, the simulation time step was 2 fs.

The entire simulated system was first equilibrated at 1 bar and 300 K, with all backbone atoms in the Mpro harmonically restrained (spring constant $k = 1 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$) for ~ 5 ns, during which the Mpro pocket was properly solvated by water molecules. In the subsequent production simulation, the restraint was removed and the Mpro was further equilibrated in the NPT ensemble. The Langevin dynamics was applied to maintain the constant temperature (300 K) in the simulated system, and the pressure was kept constant at 1 bar using the Nosé–Hoover Langevin method.¹⁶

During the 125 ns production run, the overall dimer structure was stable and the root-mean-square deviations

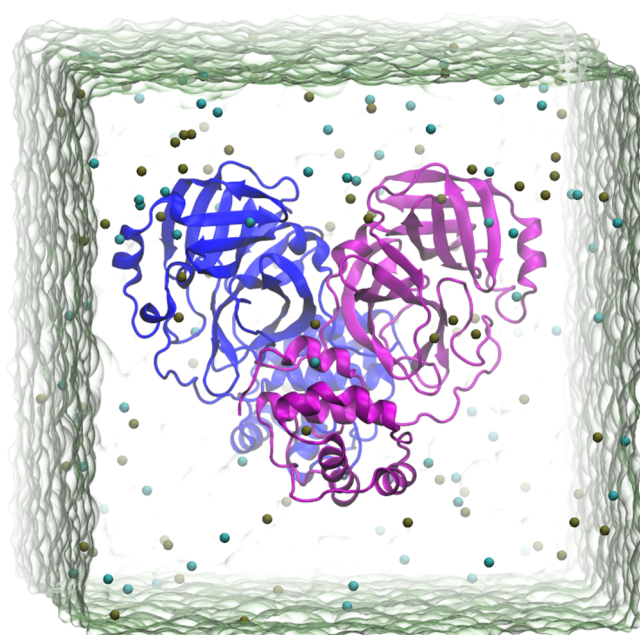


Figure 1. MD simulation system for the SARS-CoV-2's Mpro. Two monomers in the Mpro dimer (PDB entry 6LU7) are shown in cartoon representation and colored blue and purple. K^+ and Cl^- ions are shown as tan and cyan van der Waals spheres. Water is transparent.

(RMSD) of the protein's backbone calculated against the crystal (initial) one (PDB entry 6LU7) saturated at 1.7 Å after ~ 10 ns. Note that in the crystal structure there exists the N3 ligand (inside the Mpro's pocket) that is covalently linked to the Mpro, i.e., the irreversible binding. It is conceivable that the Mpro's pocket may slightly change its conformation once the N3 molecule is removed. When we were close to completing the writing of this paper, the apo structure (PDB entry 6Y2E) of the Mpro became available in the PDB. The RMSD calculated against this latest Mpro's crystal structure without any bound ligand shows a smaller mean value, indicating that without a bound ligand the equilibrated protein structure is closer to the apo structure.

Interestingly, as shown in the inset of Figure 2a, we found that during the equilibration, the positively charged N-terminal Ser1 in one monomer (of the Mpro) and the negatively charged Glu166 in the other monomer formed a stable salt bridge in the electrolyte. This observation is particularly important as the Mpro's pocket (shown as a shaded oval in the inset of Figure 2a) is right beside Glu166, indicating that this salt bridge plays a critical role in the pocket's stability. In Figure 2b, we show some small structural differences between the equilibrated Mpro's pocket (blue) and the aligned apo structure (orange), which signifies the importance of using the Mpro's structure after being equilibrated in the MD simulation for further docking studies. Using the molecular surface representation, we highlight the ligand-binding pocket of the equilibrated Mpro in Figure 2c.

For predicting the binding mode and affinity of a ligand relative to a protein, over the past few decades, many computer-aided tools and programs have been developed for both commercial and academic uses such as Glid,¹⁷ GOLD,¹⁸ MOE DOCK,¹⁹ rDOCK,²⁰ and AutoDock Vina,²¹ to name a few. In this study, for the docking part, we employed AutoDock Vina, which is a successor of the most cited

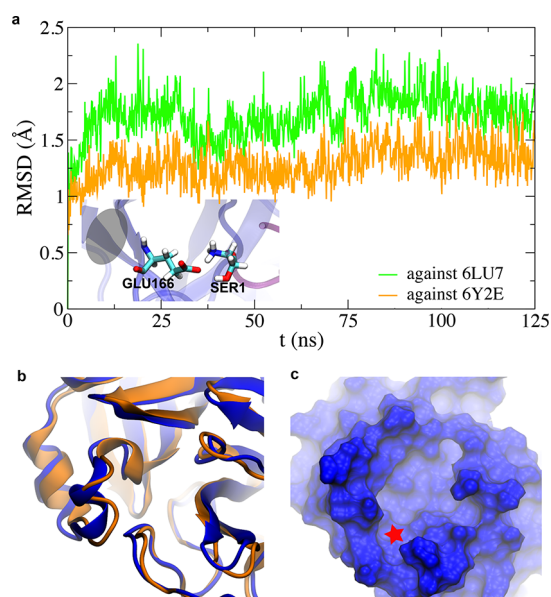


Figure 2. Molecular dynamics simulation of the SARS-CoV-2's Mpro. (a) Root-mean-square deviations (RMSD) of the simulated protein structure against the crystal (initial) structure of PDB entry 6LU7 and against the crystal structure of PDB entry 6Y2E. The inset illustrates a salt bridge formed by the negatively charged Glu166 of one monomer (blue) and the positively charged N-terminal Ser1 of the other monomer (purple) during the MD equilibration; the gray oval shows where the Mpro pocket is. (b) Cartoon representations of the equilibrated (blue) and aligned crystal Mpro's pockets without the bound N3 ligand. (c) Molecular surface representation of the equilibrated Mpro's pocket where the "anchor" site is highlighted by the star.

docking software, AutoDock,²² with significant improvement in terms of accuracy and performance. AutoDock Vina is an open-source program and was proven to have an effective scoring function.²³

We carried out the docking calculations on an IBM power node with 24 physical cores and 192 CPUs. We obtained ready-to-be-docked mol2 files for most drug molecules from the ZINC database, except for N3 and O6K whose PDB files were obtained from the RCSB PDB. For preparation of the ligands and target protein to be applied in AutoDock Vina for rigid docking, we used the scripts `prepare_ligand4.py` and `prepare_receptor4.py` provided with the AutoDock Tools^{21,22} suite to generate the corresponding input files in PDBQT format, which extends the PDB format with additional fields of partial charge and atom type. To increase the chance of finding the minimum binding energy to predict where and how a putative ligand can best bind the target protein, we set the exhaustiveness parameter of the program to 100 (the number of individual samplings/searches to find the proper pose).

Figure 3a shows the Mpro protein (equilibrated) with a ligand-binding pocket, surrounded by a rectangular box within which a ligand is docked. We first used the available crystal structure (PDB entry 6Y2F) with a bound O6K molecule (shown in Figure 3b) as a reference to calibrate the docking procedures. After searching hundreds of poses, the program yielded the best pose (shown in Figure 3c) that was very similar to that in the crystal structure (shown in Figure 3d). The obtained docking (affinity) score using the highly optimized score-function in Autodock Vina is -7.4 kcal/mol. The main difference is present in the O6K's Boc group

[containing the $-C-(CH_3)_3$ group]. In the equilibrated structure (Figure 3c), the tail resides inside the pocket region between the β -sheets (residues from Tyr161 to Asp176) and the coil (residues from Gly183 to Ala194). Hereafter, we refer to this special pocket region (also labeled with a star in Figure 2c) as the "anchor" site. However, in the crystal structure, the same Boc group is present above the β -sheet, which likely resulted from the competitive occupancy of the "anchor" site by a dimethyl sulfoxide (DMSO) molecule (Figure 3d). In Figure S1, we show from MD simulations that the Boc group of O6K occupies the "anchor" site after the removal of the DMSO molecule.

Following the same docking protocol, we evaluated 19 drugs (or compounds), most of which are currently tested in clinics for the COVID-19 disease. For these promising drugs, their antiviral mechanisms are still illusive. Given the importance of Mpro, we investigate whether the Mpro inhibition may be a part of the mechanism of action. Besides the O6K molecule shown in Figure 3b, panels a–r of Figure 4 illustrate molecular structures of chloroquine, bromhexine, favipiravir, dipyrindamole, ambroxol, hydroxychloroquine, montelukast, cinaserin, GS-441524, kaempferol, lopinavir, entecavir, umifenovir, quercetin, remdesivir, nelfinavir, curcumin, and N3, respectively. Here N3 and O6K were used in the docking studies as controls. Consistent with their presence inside the Mpro's pocket in the crystal environment, we found that two best docking scores belong to N3 (-7.1 kcal/mol) and O6K (-7.4 kcal/mol).

Remdesivir is a potent antiviral drug with an EC_{50} of 0.77 μ M toward the SARS-CoV-2 virus *in vitro*.²⁴ It was also reported to be efficacious in a few critically ill patients.²⁵ Remdesivir is a pro-drug that will hydrolyze *in vivo*. We found that both remdesivir and its metabolite (GS-441524) could form a good complex with the Mpro with affinity scores of -7.0 and -6.4 kcal/mol, respectively (Figure 4t), which could provide synergistic effects in addition to its RNA-dependent RNA polymerase (RdRp) antagonism effects. The score of remdesivir is among the best, making it a good candidate for the first-line anti-COVID-19 drug. Favipiravir is also a RdRp inhibitor developed by Fujifilm Corp. in Japan; however, this molecule is small and did not show significant binding with the Mpro in our model (Figure 4t).

Entecavir is another potential inhibitor for RdRp and had been used for treating hepatitis B virus (HBV) for years with a good efficacy and safety profile. With a structure similar to that of remdesivir's metabolite (GS-441524), entecavir is affordable and widely available and could be a good alternative as a potential inhibitor for both RdRp and Mpro. Figure 4t shows that the affinity score for entecavir is -6.4 kcal/mol, which is the same as that for GS-441524. Given its availability and promising docking score, we further carried out MD simulations to reassure the stable binding of entecavir to the Mpro's pocket (see below).

Curcumin has multifaceted function in curbing inflammation, including, IL6, TNF- α , IL-1 β , etc., and was also known to protect against liver and gastrointestinal (GI) tract damage, which is common in the COVID-19 pathological condition.²⁶ To our surprise, we found curcumin, a widely available food supplement, forms the most stable (-7.1 kcal/mol) complex with SARS-CoV-2's Mpro among the tested drugs. Its affinity score is as good as that for N3. Figure 4s shows the top pose of a curcumin molecule nicely fit inside the Mpro pocket. It is well-known that curcumin had a low bioavailability (1%) that

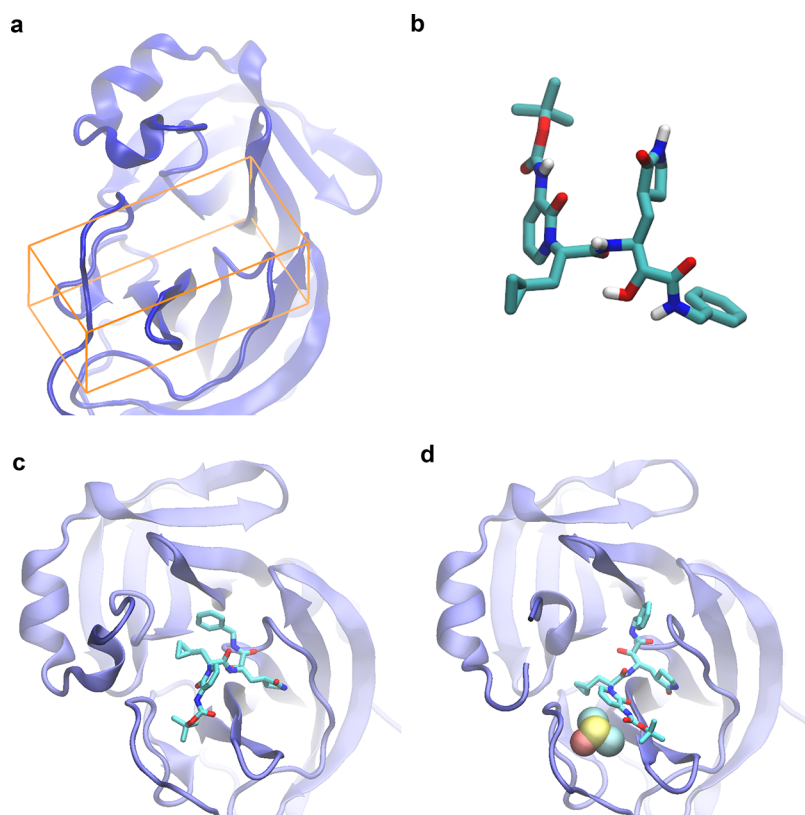


Figure 3. Docking of the O6K molecule in the pocket of the Mpro. (a) Illustration of the Mpro's pocket with a rectangular box within which a drug molecule is docked. (b) Stick representation of the O6K molecule. (c) Best docked pose of the O6K molecule in the Mpro's pocket. (d) Crystal structure (PDB entry 6Y2F) of the SARS-CoV-2's main protease with the O6K molecule and a cocrystallization agent dimethyl sulfoxide.

could hamper the utility of its treating systematic viral infection; therefore, it is important to find a viable formula that can deliver sufficient curcumin to the target organ.²⁷

However, there is some evidence that orally administered curcumin accumulates in sufficient quantity in the GI tract and liver. Besides the lung, the SARS-CoV-2 virus also infects the GI tract, causing the patients also to experience diarrhea.^{28,29} The GI tract and liver are immune privileged and may provide a shelter for the SARS-CoV-2 virus from attacks by the immune system.^{30,31} Virus from the GI tract would be shedding off even after the lung tissue is free of viral infection, which could pose a great threat to infect others in shared bathrooms or via the aerosol formed in a sewage system. Additionally, the GI system also expresses a high level of ACE2 receptors and TMPRSS2, which is critical for the SARS-CoV-2 infection. To prevent the GI system from providing a protective shelter for the SARS-CoV-2 virus, it is beneficial to take drugs (such as curcumin) to curb the virus infection of the GI system during the drug treatment regime, and even after the viral infection is tampered in the lung.

We also found quercetin and kaempferol, natural products enriched in fruits and vegetables, are capable of forming complexes with the Mpro with good binding affinities, -6.6 and -6.4 kcal/mol, respectively (Figure 4t). The bioavailability of quercetin supplements is $\sim 24\%$, compared to 52% from isoquercetin-rich onions.³² The bioavailability of kaempferol is $\sim 2\%$, many of which were metabolized in the liver. Note that both quercetin's bioavailability and kaempferol's bioavailability are better than that of curcumin.

Nelfinavir is an antiretroviral drug used in the treatment of the human immunodeficiency virus (HIV). Nelfinavir belongs

to the class of drugs known as protease inhibitors (PIs), and like other PIs, it is almost always used in combination with other antiretroviral drugs. Previously, nelfinavir has been shown to inhibit the replication of SARS-CoV.³³ Consistently, we found that the affinity score is -7.0 kcal/mol, placing it among the top candidates (Figure 4t). So far, nelfinavir has not been tested clinically for treating COVID-19 disease; however, our data strongly suggest it as a good candidate with the possible Mpro inhibition. As another HIV PI, lopinavir has an affinity score [-6.4 kcal/mol (Figure 4t)] that is lower than that for nelfinavir. According to the recent clinical trial,³⁴ lopinavir worked only modestly in the early phase of infection of SARS-CoV-2, and it is also not strong enough to support the significant efficacy at the later stage of viral infection.

Umifenovir or arbidol is an antiviral drug available in Russian and China and had been used clinically in treating COVID-19 in China. Umifenovir inhibits SARS-CoV-2 *in vitro* with a reported IC_{50} of $30 \mu M$.³⁵ It has also demonstrated positive results as a postexposure prophylaxis (PEP) of COVID-19 transmission.³⁶ We obtained moderate Mpro binding in our model with a score of -6.5 kcal/mol (Figure 4t).

Montelukast is a leukotriene inhibitor used to treat allergies and prevent asthma attacks. It was reported anecdotally to inhibit the Mpro, and we validated its binding with the Mpro with a moderate score of -6.2 kcal/mol (Figure 4t). Montelukast could offer other anti-inflammation benefits other than the potential Mpro inhibition.

Bromhexine³⁷ and ambroxol³⁸ are over the counter expectorant drugs and are used in assisting the treatment of COVID-19 in China. We found that they may exhibit weak

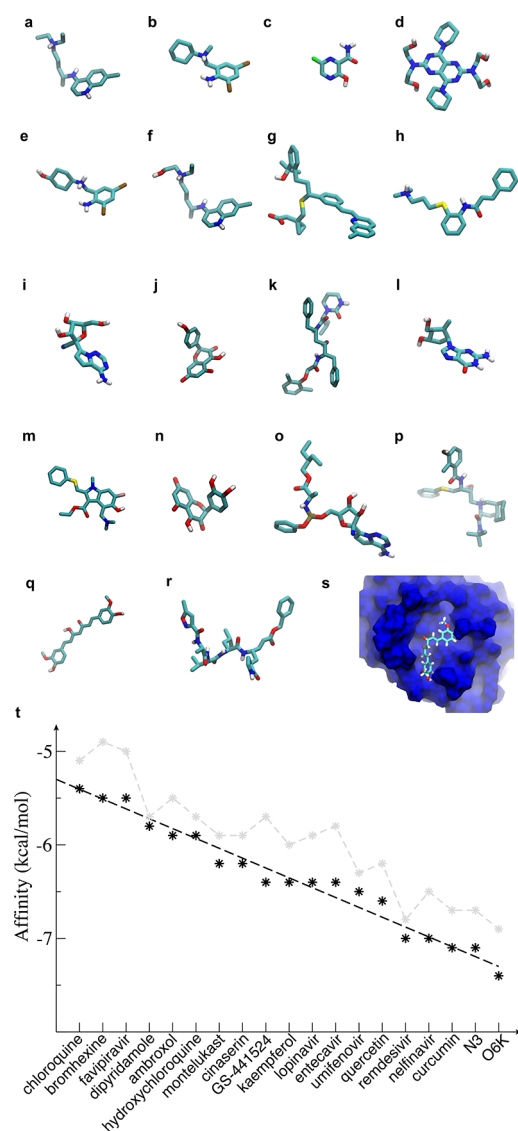


Figure 4. Docking of various clinically tried drugs in the pocket of the Mpro: (a) chloroquine, (b) bromhexine, (c) favipiravir, (d) dipyridamole, (e) ambroxol, (f) hydroxychloroquine, (g) montelukast, (h) cinaserin, (i) GS-441524, (j) kaempferol, (k) lopinavir, (l) entecavir, (m) umifenovir, (n) quercetin, (o) remdesivir, (p) nelfinavir, (q) curcumin, and (r) N3. (s) Top pose of curcumin in the Mpro pocket. (t) Best (black) and 10th best (gray) binding affinities of tested drug molecules docked in the Mpro pocket.

Mpro antagonism effects according to docking affinities (Figure 4t). Nevertheless, expectorant drugs normally are highly enriched in lung tissues, which can promote the inhibition of Mpro's activities.

Chloroquine was found to inhibit SARS-CoV-2 *in vitro*²⁹ in clinical tests. Despite very similar chemical properties, the efficacy of hydroxychloroquine (HCQ) is 7 times stronger than that of chloroquine (CQ) *in vitro* (EC_{50} 0.7 μ M vs 5.4 μ M in the Vero cell infection model).³⁹ The potency discrepancy calls for another antiviral mechanism besides the existing hypothesis of HCQ/CQ functioning by neutralizing the endosomal pH and inhibiting cathepsin L,⁴⁰ which would predict a similar efficacy if not the same, with such minor structural differences between HCQ and CQ.

In support of other unknown targets by HCQ and CQ, the presence of extracellular proteases like TMPRSS2 and Furin

facilitates efficient virus entry on the plasma membrane directly and bypass on the endocytosis pathway.⁴¹ The fact that HCQ/CQ still poses strong inhibition of virus clinically also suggests HCQ/CQ inhibited an intracellular common pathway such as the Mpro. Meanwhile, HCQ/CQ can accumulate in the lung with a concentration that is 200–700-fold higher than that of the serum,⁴² which may suffice for Mpro inhibition with suboptimal affinity scores of -5.9 kcal/mol for HCQ and -5.0 kcal/mol for CQ (see Figure 4t).

Dipyridamole exhibits a broad spectrum of antiviral activities and has been used as an effective anti-COVID-19 drug clinically.⁴³ The antiviral mechanism of dipyridamole involves the inhibition of the Mpro as determined by the surface plasmon resonance (SPR) assay *in vitro*.⁴³ However, the affinity score is -5.8 kcal/mol (Figure 4t). Additionally, cinaserin was predicted to inhibit the Mpro of both SARS-CoV and SARS-CoV-2; however, we found its binding affinity with Mpro is not very high, in line with a 120 μ M IC_{50} for the viral inhibition.⁶ The analogue of cinaserin (cmpd-26) performs much better in wet-lab experiments with an IC_{50} of 1.06 μ M for SARS-CoV Mpro protease;⁴⁴ however, the score of cmpd-26 is not much improved for the SARS-CoV-2 Mpro. These discrepancies may be due to the rough estimations of the binding affinity from the docking study, which suggests the need for a more accurate model such as the MD one that includes the impact of water (e.g., desolvation energy).

To verify the docking results, as an example, we further performed the MD simulation to investigate the stability of entecavir's pose with the best affinity score inside the Mpro pocket (Figure 5a). The force field parameters for entecavir were obtained from SwissParam.⁴⁵ We applied the same MD simulation protocol used for the apo Mpro equilibration (Figures 1 and 2). To highlight the positions of the entecavir molecule inside the Mpro pocket, we overlapped the center of mass (COM) of entecavir during the entire MD simulation in Figure 5b. Notably, there are two clusters of COMs, representing the start and final locations of the entecavir molecule. The entire process for entecavir's repositioning its pose inside the Mpro pocket is illustrated in the movie in the Supporting Information.

Figure 5c shows the initial pose of the entecavir molecule in the MD simulation, which is also the best pose from the docking study. Inside the Mpro pocket, the entecavir molecule formed hydrogen bonds with residues Thr26, His41, and Gly143. Overall, entecavir is smaller than the Mpro's pocket, and the "anchor" site (the orange oval in Figure 5c) mentioned above is not occupied. After entecavir has drifted to the new location inside the Mpro, it coordinates with His41, Gln189, and Glu166 through hydrogen bonds. Interestingly, the five-membered ring in the entecavir molecule entered the "anchor" site (see the orange oval in Figure 5d), driven by the hydrophobic interaction. Note that the new pose of entecavir found in the MD simulation is similar to the eighth pose with a slightly higher score (-5.9 kcal/mol) discovered in the docking study. This highlights that when including the solvent effect the affinity scores listed in Figure 4t can be different ($\sim 10\%$), which can somewhat affect the rank order.

To further quantitatively demonstrate the dynamic drifting process of entecavir, we calculated the time-dependent contact area between the entecavir molecule and the Mpro pocket. On the basis of the concept of the solvent-accessible surface area, we calculated the surface areas of the Mpro pocket (S_M), the entecavir molecule (S_E), and their complex (S_T). Thus, we

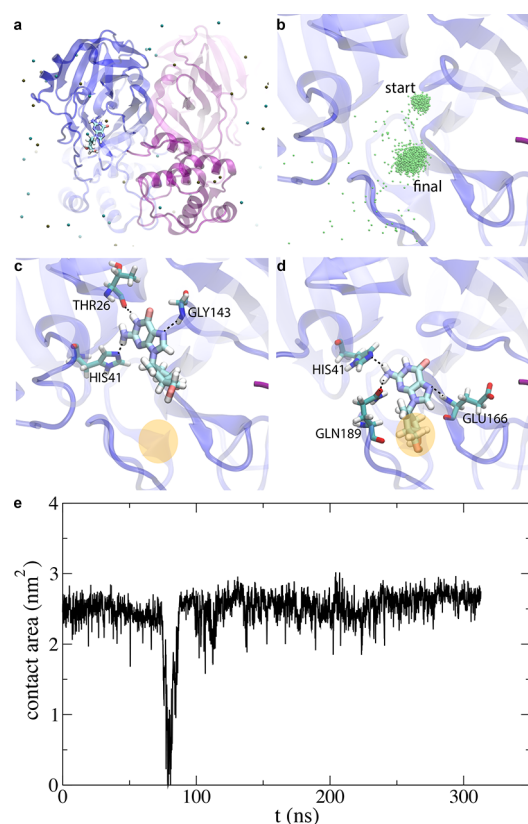


Figure 5. Molecular dynamics simulation of entecavir in the Mpro pocket. (a) Enlarged view of the MD simulation system for the Mpro with a bound entecavir molecule. Water is not shown for the sake of clarity. (b) Centers of mass of the entecavir molecule during the entire MD simulation. (c) Entecavir's pose at the beginning of the MD simulation, obtained from the docking study. (d) Entecavir's pose after the MD equilibration. (e) Time dependency of the contact area between the entecavir molecule and the Mpro pocket.

define the contact area between the entecavir molecule and the Mpro pocket as $(S_M + S_E - S_T)/2$. Figure 5e shows the calculated contact area during the >300 ns MD simulation. In the first ~75 ns, the entecavir molecule stayed in its initial location with an average contact area of ~2.4 nm². After that, it drifted away from its initial location and move into water above the Mpro pocket. The contact area decreased to approximately zero. Without drifting further into water, the entecavir molecule reentered the Mpro pocket, and the contact area increased to ~2.6 nm². During the rest of the MD simulation (~250 ns), the entecavir molecule remained stable in the new location. In an independent MD simulation, the entecavir molecule left its initial location within tens of nanoseconds and drifted into the water environment, which reassures that the entecavir's initial pose is less stable than the final pose revealed in the MD.

Among all of the drug ligands studied here, nelfinavir appears to be highly promising as the Mpro's inhibitor, suggested by our docking study (Figure 4). Thus, we carried out two independent MD simulations (Sim-1 and Sim-2) for nelfinavir in Mpro's pocket (Figure 6a) to corroborate our prediction. From the two independent MD simulations, we found that the nelfinavir molecule was stably bound inside the Mpro's pocket, as shown by the overlapped COMs of nelfinavir (Figure 6b for Sim-1 and Figure S2 for Sim-2). The stability was contributed by the internal hydrogen bond

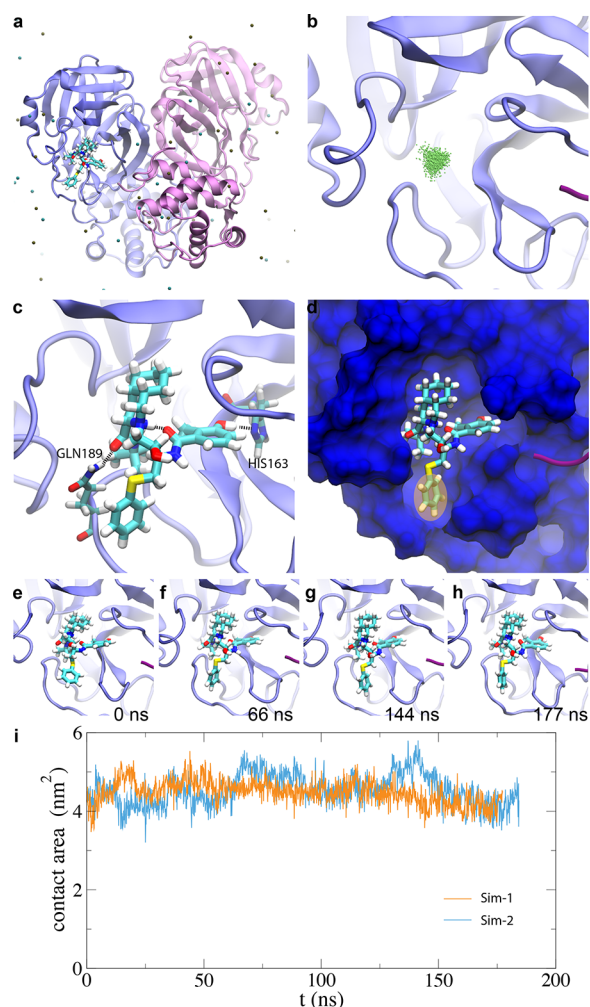


Figure 6. Molecular dynamics simulation of nelfinavir in the Mpro pocket. (a) Enlarged view of the MD simulation system for the Mpro with a bound nelfinavir molecule. Water is not shown for the sake of clarity. (b) Centers of mass of the nelfinavir molecule during the entire MD simulation (Sim-1). (c) Representative pose of nelfinavir in the Mpro pocket. (d) Nelfinavir's pose in the Mpro (in the molecular surface representation). (e–h) Snapshots of nelfinavir's conformations in the Mpro's pocket from Sim-1. (i) Time dependency of the contact area between the nelfinavir molecule and the Mpro pocket, from two independent MD simulations (Sim-1 and Sim-2).

between the -NH and -C=O groups (Figure 6c), two hydrogen bonds between nelfinavir and residues His163 and Gln189 in Mpro (Figure 6c), and the significant amount of hydrophobic interactions between nelfinavir and Mpro. Overall, the entire molecule was nicely fit inside the Mpro's pocket, with its benzene group located at the "anchor" site (Figure 6d). We further highlight the stable binding of nelfinavir with snapshots at simulation times of 0, 66, 144, and 177 ns (Figure 6e,f) and with a movie in the Supporting Information showing the simulation trajectory (Sim-1). Therefore, the predicted pose of nelfinavir in the Mpro's pocket, from docking, was confirmed by MD simulations. Compared with entecavir, nelfinavir almost occupies the whole Mpro's pocket and has a much larger contact area (~4.5 nm²) with the Mpro, as shown in the two independent MD simulations (Sim-1 and Sim-2) in Figure 6i.

Through both the docking and MD studies of the 19 drug molecules, we found that the "anchor" site in the Mpro pocket

plays an important role in stabilizing a bound ligand. For example, in the docking study, one end of the curcumin molecule was seen to insert inside the “anchor” site (Figure 4s). Likewise, in MD simulations, the entecavir molecule found a more stable binding pose with its five-membered ring inside the “anchor” site and the nelfinavir molecule was stably bound inside the Mpro’s pocket with its benzene group residing inside the “anchor” site. In the crystal structure (PDB entry 6LU7), the tail of the N3 molecule (Figure 4r) also occupies the “anchor” site of the Mpro, evidencing the crucial role of the “anchor” site for ligand binding.

In a recent experimental study,⁵ it was found that the bulky Boc group in the O6K molecule (Figure 3b) is essential for its binding to the Mpro pocket, without which the ligand becomes inactive. As shown in Figure 3c, our docking result verifies that the Boc group indeed occupies the “anchor” site, further validating the binding mechanisms discovered in this work for stabilizing the ligand inside the Mpro’s pocket.

Additionally, we point out that in the crystal structure (PDB entry 6Y2F), there is a DMSO molecule binding inside the “anchor” site (Figure 3d), which may function to impede the binding of potential inhibitors. It is well-known that DMSO is used widely to dissolve compounds for biological assays. For example, dipyrindamole had a low solubility and inhibited Mpro in an SPR assay with a 34 μ M binding affinity. However, the *in vitro* viral cell-based assay showed that the EC₅₀ of dipyrindamole can be as low as 0.1 μ M. The SPR assay utilized purified Mpro proteins exposed directly to a compound solution containing 0.5% DMSO; the latter could compete with dipyrindamole for binding with Mpro and results in a false negative outcome.

In summary, we have used *in silico* methods, including docking and MD simulation, to investigate the stability of various ligands bound inside the Mpro pocket. Besides the binding energy, we have also considered comprehensively the bioavailabilities and half-lives of the drugs, which translate into effective drug concentrations around the targets. We found that the docking affinity scores of several molecules (such as nelfinavir and entecavir) are very close to those of the ligands found experimentally (N3 and O6K), and their binding stabilities with the Mpro were verified in MD simulations. Indeed, during the review of this work, a preprint published in bioRxiv shows that experimentally nelfinavir can strongly inhibit the Mpro *in vitro*.⁴⁶ However, we note that there are potential limitations of our methods, such as the approximated scoring functions in the docking method and classic force fields in the MD method. Therefore, the rank order provided in Figure 4t is by no means intended to promote certain drug molecules. Instead, by examining almost 100 different poses of each docked molecule, we realized the importance of the “anchor” site in the Mpro pocket. Besides the hint from docking studies, we confirmed in MD simulation that by occupying the “anchor” site, a ligand can reside inside the Mpro pocket more stably via hydrophobic interactions. More importantly, the binding mechanism revealed in this work is supported by evidence observed in experimental studies, including the recently published one (during the review of this work) showing the bindings of compounds 11a and 11b with a hydrophobic group in the “anchor” site.⁴⁷ Therefore, we conclude that a drug molecule can be more potent with a hydrophobic bulky group occupying the “anchor” site. Given the bottleneck in the current testing capacity and throughput of the Mpro protease assay, we expect that these *in silico* results

and predictions could help the wet-lab groups to prioritize their testing and shed light on the future discovery of highly potent inhibitors targeting the SARS-CoV-2’s Mpro.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpclett.0c00994>.

MD simulation of O6K in the Mpro’s pocket (Figure S1) and second MD simulation (Sim-2) for nelfinavir in the Mpro’s pocket (Figure S2) (PDF)

Trajectories of the MD simulation for the entecavir–Mpro complex (MPG)

Trajectories of the MD simulation for the nelfinavir–Mpro complex (MPG)

■ AUTHOR INFORMATION

Corresponding Author

Binquan Luan – Computational Biological Center, IBM Thomas J. Watson Research, Yorktown Heights, New York 10598, United States; orcid.org/0000-0002-9414-5379; Email: bluan@us.ibm.com

Authors

Tien Huynh – Computational Biological Center, IBM Thomas J. Watson Research, Yorktown Heights, New York 10598, United States

Haoran Wang – Neoland Biosciences, Medford, Massachusetts 02155, United States

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jpclett.0c00994>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

B.L. gratefully acknowledges the computing resource and the financial support from the IBM Bluegene Science Program.

■ REFERENCES

- (1) Cui, J.; Li, F.; Shi, Z.-L. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **2019**, *17*, 181–192.
- (2) Kuiken, T.; Fouchier, R. A.; Schutten, M.; Rimmelzwaan, G. F.; Van Amerongen, G.; van Riel, D.; Laman, J. D.; de Jong, T.; van Doornum, G.; Lim, W.; et al. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet* **2003**, *362*, 263–270.
- (3) Zaki, A. M.; Van Boheemen, S.; Bestebroer, T. M.; Osterhaus, A. D.; Fouchier, R. A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* **2012**, *367*, 1814–1820.
- (4) Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **2003**, *300*, 1763–1767.
- (5) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerharing, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* **2020**, eabb3405.
- (6) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C. Structure of Mpro from COVID-19 virus and discovery of its inhibitors. *bioRxiv* **2020**, DOI: 10.1101/2020.02.26.964882.
- (7) Azad, G. K.; Tomar, R. S. Ebselen, a promising antioxidant drug: mechanisms of action and targets of biological pathways. *Mol. Biol. Rep.* **2014**, *41*, 4865–4879.

- (8) Karplus, M.; McCammon, J. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646.
- (9) Luan, B.; Xu, G.; Feng, M.; Cong, L.; Zhou, R. Combined Computational–Experimental Approach to Explore the Molecular Mechanism of SaCas9 with a Broadened DNA Targeting Range. *J. Am. Chem. Soc.* **2019**, *141*, 6545–6552.
- (10) Phillips, J. C.; et al. *J. Comput. Chem.* **2005**, *26*, 1781.
- (11) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (12) Neria, E.; Fischer, S.; Karplus, M. Simulation of Activation Free Energies in Molecular Systems. *J. Chem. Phys.* **1996**, *105*, 1902–1921.
- (13) Beglov, D.; Roux, B. Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. *J. Chem. Phys.* **1994**, *100*, 9050–9063.
- (14) Huang, J.; MacKerell, A. D., Jr CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34*, 2135–2145.
- (15) Miyamoto, S.; Kollman, P. A. SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Molecules. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (16) Martinetz, T.; Schulten, K. Topology Representing Networks. *Neur. Netw.* **1994**, *7*, 507–522.
- (17) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (18) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (19) Corbeil, C. R.; Williams, C. I.; Labute, P. Variability in docking success rates due to dataset preparation. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 775–786.
- (20) Ruiz-Carmona, S.; Alvarez-Garcia, D.; Foloppe, N.; Garmendia-Doval, A. B.; Juhos, S.; Schmidtke, P.; Barril, X.; Hubbard, R. E.; Morley, S. D. rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Comput. Biol.* **2014**, *10*, No. e1003571.
- (21) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2009**, *31*, 455–461.
- (22) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.
- (23) Wang, Z.; Sun, H.; Yao, X.; Li, D.; Xu, L.; Li, Y.; Tian, S.; Hou, T. Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power. *Phys. Chem. Chem. Phys.* **2016**, *18*, 12964–12975.
- (24) Wang, M.; Cao, R.; Zhang, L.; Yang, X.; Liu, J.; Xu, M.; Shi, Z.; Hu, Z.; Zhong, W.; Xiao, G. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **2020**, *30*, 269–271.
- (25) Holshue, M. L.; DeBolt, C.; Lindquist, S.; Lofy, K. H.; Wiesman, J.; Bruce, H.; Spitters, C.; Ericson, K.; Wilkerson, S.; Tural, A.; et al. First case of 2019 novel coronavirus in the United States. *N. Engl. J. Med.* **2020**, *382*, 929.
- (26) Hewlings, S. J.; Kalman, D. S. Curcumin: a review of its' effects on human health. *Foods* **2017**, *6*, 92.
- (27) Prasad, S.; Tyagi, A. K.; Aggarwal, B. B. Recent developments in delivery, bioavailability, absorption and metabolism of curcumin: the golden pigment from golden spice. *Cancer research and treatment: official journal of Korean Cancer Association* **2014**, *46*, 2.
- (28) Zhang, W.; Du, R.-H.; Li, B.; Zheng, X.-S.; Yang, X.-L.; Hu, B.; Wang, Y.-Y.; Xiao, G.-F.; Yan, B.; Shi, Z.-L.; et al. Molecular and serological investigation of 2019-nCoV infected patients: implication of multiple shedding routes. *Emerging Microbes Infect.* **2020**, *9*, 386–389.
- (29) Wang, W.; Xu, Y.; Gao, R.; Lu, R.; Han, K.; Wu, G.; Tan, W. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *Jama, J. Am. Med. Assoc.* **2020**, DOI: 10.1001/jama.2020.3786.
- (30) Wohlleber, D.; Knolle, P. A. *Infection, Immune Homeostasis and Immune Privilege*; Springer, 2012; pp 93–106.
- (31) Neil, J. A.; Cadwell, K. The intestinal virome and immunity. *J. Immunol.* **2018**, *201*, 1615–1624.
- (32) Kaşıkçı, M. B.; Bağdatlıoğlu, N. Bioavailability of quercetin. *Current research in nutrition and food science journal* **2016**, *4*, 146–151.
- (33) Yamamoto, N.; Yang, R.; Yoshinaka, Y.; Amari, S.; Nakano, T.; Cinatl, J.; Rabenau, H.; Doerr, H. W.; Hunsmann, G.; Otaka, A.; et al. HIV protease inhibitor nelfinavir inhibits replication of SARS-associated coronavirus. *Biochem. Biophys. Res. Commun.* **2004**, *318*, 719–725.
- (34) Cao, B.; Wang, Y.; Wen, D.; Liu, W.; Wang, J.; Fan, G.; Ruan, L.; Song, B.; Cai, Y.; Wei, M.; et al. A trial of lopinavir–ritonavir in adults hospitalized with severe Covid-19. *N. Engl. J. Med.* **2020**, *382*, 1787.
- (35) Lu, H. Drug treatment options for the 2019-new coronavirus (2019-nCoV). *BioSci. Trends* **2020**, *14*, 69.
- (36) Zhang, J.; Wang, W.; Peng, B.; Peng, W.; Zhang, Y.; Wang, Y.; Wan, Y.; Chang, J.; Mao, L.; Miao, X. Potential of Arbidol for Post-exposure Prophylaxis of COVID-19 Transmission. *Chinaxiv*, 2020.
- (37) Yang, Y.; Shen, H.; Zhaoshan, M.; Wang, D.; Wang, H. Potential mechanism of bromhexine in the treatment of COVID-2019. *Chinese Pharmaceutical Journal*, 2020.
- (38) Li, M.; Wang, J. Prospect of ambroxol in the treatment of COVID-2019. *Chinese Journal of Clinical Pharmacology*, 2020.
- (39) Yao, X.; Ye, F.; Zhang, M.; Cui, C.; Huang, B.; Niu, P.; Liu, X.; Zhao, L.; Dong, E.; Song, C.; et al. In vitro antiviral activity and projection of optimized dosing design of hydroxychloroquine for the treatment of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *Clin. Infect. Dis.* **2020**, DOI: 10.1093/cid/ciaa237.
- (40) Al-Bari, M. A. A. Targeting endosomal acidification by chloroquine analogs as a promising strategy for the treatment of emerging viral diseases. *Pharmacol. Res. Perspect.* **2017**, *5*, e00293.
- (41) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A.; et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **2020**, *181*, 271.
- (42) McChesney, E. W.; Banks, W. F., Jr; Fabian, R. J. Tissue distribution of chloroquine, hydroxychloroquine, and desethylchloroquine in the rat. *Toxicol. Appl. Pharmacol.* **1967**, *10*, 501–513.
- (43) Liu, X.; Li, Z.; Liu, S.; Chen, Z.; Zhao, Z.; Huang, Y.-y.; Zhang, Q.; Wang, J.; Shi, Y.; Xu, Y. Therapeutic effects of dipyrindamole on COVID-19 patients with coagulation dysfunction. *MedRxiv* **2020**, DOI: 10.1101/2020.02.27.20027557.
- (44) Yang, Q.; Chen, L.; He, X.; Gao, Z.; Shen, X.; Bai, D. Design and synthesis of cinanserin analogs as severe acute respiratory syndrome coronavirus 3CL protease inhibitors. *Chem. Pharm. Bull.* **2008**, *56*, 1400–1405.
- (45) Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: a fast force field generation tool for small organic molecules. *J. Comput. Chem.* **2011**, *32*, 2359–2368.
- (46) Ohashi, H.; Watashi, K.; Saso, W.; Shionoya, K.; Iwanami, S.; Hirokawa, T.; Shirai, T.; Kanaya, S.; Ito, Y.; Kim, K. S. Multidrug treatment with nelfinavir and cepharanthine against COVID-19. *bioRxiv* **2020**, DOI: 10.1101/2020.04.14.039925.
- (47) Dai, W.; Zhang, B.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Liu, F.; Li, C.; Li, Y. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **2020**, eabb4489.